Single View 3D Reconstruction under an Uncalibrated Camera and an Unknown Mirror Sphere

Kai Han, Kwan-Yee K. Wong, Xiao Tan
The University of Hong Kong,* Hong Kong
{khan, kykwong}@cs.hku.hk, tanxchong@gmail.com

Abstract

In this paper, we develop a novel self-calibration method for single view 3D reconstruction using a mirror sphere. Unlike other mirror sphere based reconstruction methods, our method needs neither the intrinsic parameters of the camera, nor the position and radius of the sphere be known. Based on eigen decomposition of the matrix representing the conic image of the sphere and enforcing a repeated eignvalue constraint, we derive an analytical solution for recovering the focal length of the camera given its principal point. We then introduce a robust algorithm for estimating both the principal point and the focal length of the camera by minimizing the differences between focal lengths estimated from multiple images of the sphere. We also present a novel approach for estimating both the principal point and focal length of the camera in the case of just one single image of the sphere. With the estimated camera intrinsic parameters, the position(s) of the sphere can be readily retrieved from the eigen decomposition(s) and a scaled 3D reconstruction follows. Experimental results on both synthetic and real data are presented, which demonstrate the feasibility and accuracy of our approach.

1. Introduction

3D reconstruction has always been a hot topic in the field of computer vision. Tremendous efforts have been devoted to this problem in the past decades. In particular, multi-view stereo (MVS) has been one of the most popular and successful approaches in solving this problem, and numerous state-of-the-art MVS methods have been proposed (e.g., [18]). Single view approach, on the other hand, has received much less attention compared with MVS. Despite its potential, single view approach is relatively less studied in the literature.

The working principle of most single view methods is based on observing multiple light paths to the same scene point. This is often accomplished by introducing one or more mirrors into the scene and observing the reflection(s) of the scene on the mirror(s). Both planar and spherical mirrors have been employed by such methods. Planar mirrors do not introduce any distortion in the reflected images, but provide only a very limited field of view (FOV) (e.g., [9, 10, 11]). Spherical mirrors, on the other hand, can provide a much wider FOV for 3D reconstruction (e.g., [8, 17, 20, 21]). However, distortions do exist in the reflected images, and these make the correspondence problem more difficult.

Most of the existing mirror based methods for single view 3D reconstruction require a fully calibrated setup, in which the intrinsic parameters of both the camera and mirrors, as well as the positions and orientations of the mirrors (relative to the camera) are assumed to be known. This often requires tedious calibration that hinders the application of these methods. In [7], Chen *et al.* showed that, with an internally calibrated camera, it is possible to recover the position of a mirror sphere from its image up to a scale determined by its radius. This allows a 3D reconstruction up to an unknown scale under an unknown radius of the mirror sphere.

In this paper, we revisit the problem studied in [7], which is single view 3D reconstruction using an unknown mirror sphere, but with an uncalibrated camera. It is well known that, under perspective projection, the image of a sphere would be a conic [12]. Based on eigen decomposition of the matrix representing the conic image and enforcing a repeated eignvalue constraint, we derive an analytical solution for recovering the focal length of the camera given its principal point. Based on this analytical solution, we develop two robust algorithms for estimating both the principal point and focal length of the camera from multiple images as well as from just one single image of the mirror sphere. With the estimated camera intrinsic parameters, the position(s) of the sphere can be readily retrieved from the eigen decomposition(s) as in [7], and a scaled 3D reconstruction follows. The key contributions of this work include

- To the best of our knowledge, the first single view 3D reconstruction method that works under an uncalibrated camera and an unknown mirror sphere (i.e., one with both its position and radius being unknown).

- An analytical solution for recovering the focal length of a camera from an image of an unknown sphere given the principal point of the camera.

- A robust method for estimating both the principal point and focal length of a camera from multiple images of an unknown sphere placed at different positions.

- A novel method for estimating both the principal point and focal length of a camera from just one single image of an unknown sphere.

The rest of this paper is organized as follows. Section 2 gives a brief literature review. Section 3 provides the theoretical background of this work. The proposed algorithms for estimating both the principal point and focal length of a camera are introduced in Section 4, followed by a brief description of the reconstruction method in Section 5. Experimental results are presented in Section 6. Section 7 discusses and concludes this work.

2. Related Work

Imaging systems consisting of a camera observing one or more mirrors are referred to as a catadioptric imaging system. They have many applications in both computer graphics and comuputer vision [27], including panoramic imaging [34], stereo vision [8, 9, 11, 19, 23], light field imaging [33], recognition [28], etc. Both planar mirrors and spherical mirrors are commonly used in constructing a catadioptric imaging system.

A planar mirror provides a very cheap way of constructing a new viewpoint, and is the simplest device for building a stereo vision system from a single camera (e.g., [10]). It is useful in applications like 3D reconstruction [13, 39] and light field imaging [16]. In [22], Mitsumoto et al. described the single planar mirror geometric constraints for 3D reconstruction. They showed that it is possible to recover a large coverage of an object by moving the planar mirror or by placing multiple planar mirrors around the object. In [11], Gluckman and Nayar used multiple planar mirrors to build a complex imaging geometry. A planar mirror has the advantage of not introducing any distortion in the reflected image. However, its small FOV greatly hinders its use in practice.

A spherical mirror can provide a much wider FOV than a planar mirror, and may even reflect the entire surrounding environment. This makes it more commonly used in catadioptric systems. Existing methods often assume an internally calibrated camera, and consider only extrinsic calibration (e.g., [1, 2, 25, 31]). In [23], Nayar presented the

sphereo system for scene depth recovery. It consists of a calibrated camera looking at two specular spheres, with both the radii and positions of the spheres being known a priori. In [26], Powell et al. recovered light source positions from specular highlights observed on surfaces of spherical mirrors. In [20], Lanman et al. built a catadioptric system using a perspective camera and many identical spherical mirrors. They recovered the parameters of the spherical mirrors by a tailored calibration method. In [19], Kanbara et al. attached a color marker around the camera lens to estimate the visual ray passing through a spherical mirror with a known radius. In [30, 35], Wong et al. estimated the camera poses and light source directions from a sphere with unknown radius and position using specular highlights observed on the sphere and the silhouettes of the sphere. In [24] and [29], the corneas were approximated as spheres, and exploited for reconstructing a display from its reflections on the corneas. In [7], Chen et al. introduced a method to reconstruct a 3D object using a moving spherical mirror. They showed that without knowing the radius of the spherical mirror, a 3D surface can be recovered up to an unknown scale. Note that all the aforementioned methods assume an internally calibrated camera or a known sphere.

Other than using mirrors to build a catadioptric system, there is also much effort aiming at recovering the mirror surface itself from reflections observed on its surface (e.g., [4, 14, 15]). These, however, are out of the scope of this paper. Besides, without considering the catadioptric system, camera calibration has been explored using diffuse spheres (e.g., [3, 32, 36, 37, 38]), while these methods normally require multiple images and multiple spheres.

3. Theoretical Background

Without loss of generality, consider a pinhole camera with its optical centre locates at the origin $\mathbf{O}$ of a world coordinate system, and a sphere of radius r_s with its centre locates at a distance d from $\mathbf{O}$ along the positive Z-axis of the world coordinate system (see Fig. 1). The projection matrix of this camera can be written as $\mathbf{P} = \mathbf{K}[\mathbf{R} \ \mathbf{0}]$, where $\mathbf{K}$ is the 3×3 camera calibration matrix composed of the camera intrinsic parameters, and $\mathbf{R}$ and $\mathbf{0}$ are the 3×3 rotation matrix and the translation vector $[0 \ 0 \ 0]^{\mathrm{T}}$, respectively, which define the rigid body transformation of the camera with respect to the world coordinate system. The sphere will project onto the image as a conic. This conic can be represented by a 3×3 symmetric matrix $\mathbf{C}$ such that each point $\mathbf{x}$ (in homogeneous pixel coordinates) lying on this conic satisfies $\mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} = 0$. Such a conic can be easily obtained from the image by applying a robust conic fitting algorithm [12].

By removing the effect of $\mathbf{K}$, the conic $\mathbf{C}$ will transform into

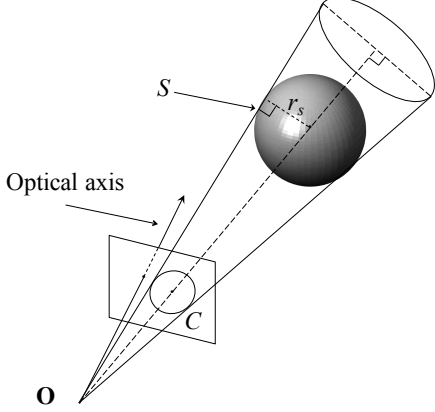$$\hat{\mathbf{C}} = \mathbf{K}^{\mathrm{T}}\mathbf{C}\mathbf{K} \qquad (1)$$

Figure 1. A perspective camera located at $\mathbf{O}$ observes a sphere S of radius r_s. The image of the sphere S is a conic C, which can be represented by a a 3×3 symmetric matrix $\mathbf{C}$. $\mathbf{O}$ and C define a right circular cone tangent to S. The axis of this cone pierces the center of S. This axis is also the Z-axis of the world coordinate, and the distance between $\mathbf{O}$ and the center of S is d.

which is a conic expressed in image plane coordinates. It has been shown in [35] that, by eign decomposition, $\hat{\mathbf{C}}$ can be factorized into

$$\hat{\mathbf{C}} = \mathbf{R}\mathbf{D}\mathbf{R}^{\mathrm{T}}, \tag{2}$$

where $\mathbf{R}$ is the rotation matrix of the camera, and $\mathbf{D} = diag(\lambda_1, \lambda_1, \lambda_2)$ is a diagonal matrix composed of the eigenvalues of $\hat{\mathbf{C}}$. Note that $\mathbf{D}$ represents a circle of radius $r_c = \sqrt{-\frac{\lambda_1}{\lambda_2}}$ centres at the image plane origin. It corresponds to the image of the sphere when the camera has its optical axis aligned with the Z-axis of the world coordinate system (i.e., when $\mathbf{R} = \mathbf{I}$). The sphere centre, expressed in camera coordinates, can be recovered as $d\mathbf{r}_3$, where $\mathbf{r}_3$ is the third column of $\mathbf{R}$ and $d = r_s \sqrt{\frac{1+r_c^2}{r_c}}$.

With a calibrated camera and a sphere with known radius, the position of the sphere can be uniquely recovered from its image. An object can be reconstructed from its reflections on the sphere placed at two distinct positions. When the radius of the sphere is not known, the position of the sphere can still be recovered up to an unknown scale determined by this unknown radius, and this results in a reconstruction up to the same unknown scale.

4. Estimating Camera Intrinsic Parameters

In this section, we first derive an analytical solution for recovering the focal length of a camera from an image of a sphere under a known principal point of the camera. Based on this analytical solution, we introduce two robust algorithms for estimating both the principal point and focal length of the camera from multiple images of the sphere as well as from just one single image of the sphere.

4.1. Focal Length

Assume the camera has unit aspect ratio, and let f and (u_0, v_0) be its focal length and principal point respectively. The camera calibration matrix $\mathbf{K}$ can be factorized into

$$\mathbf{K} = \mathbf{T}\mathbf{F}, \tag{3}$$

where

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & u_0 \\ 0 & 1 & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{F} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

Substituting (3) into (1) gives

$$\hat{\mathbf{C}} = \mathbf{F}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{C}\mathbf{T}\mathbf{F}. \tag{5}$$

Now suppose the principal point of the camera is known (e.g., by assuming the principal point is at the image centre). The effect of $\mathbf{T}$ can be removed by translating all the points by $(-u_0, -v_0)$. After the translation, the conic $\mathbf{C}$ will transform into

$$\bar{\mathbf{C}} = \mathbf{T}^{\mathrm{T}}\mathbf{C}\mathbf{T} \tag{6}$$

which is a conic having the same shape as $\mathbf{C}$ (see Fig. 2).
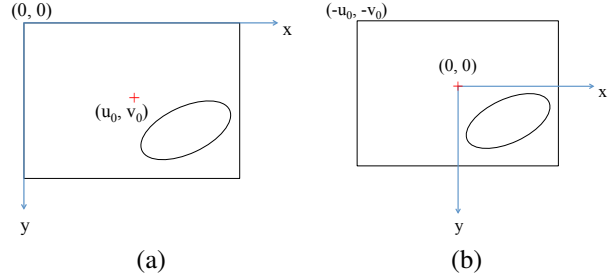

Figure 2. (a) $\mathbf{C}$: before removing the effect of $\mathbf{T}$. (b) $\bar{\mathbf{C}}$: after removing the effect of $\mathbf{T}$.

Substituting (6) into (5) gives

$$\hat{\mathbf{C}} = \mathbf{F}^{\mathrm{T}}\bar{\mathbf{C}}\mathbf{F}. \tag{7}$$

Now let

$$\hat{\mathbf{C}} = \begin{bmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} & \hat{\omega}_{13} \\ \hat{\omega}_{21} & \hat{\omega}_{22} & \hat{\omega}_{23} \\ \hat{\omega}_{31} & \hat{\omega}_{32} & \hat{\omega}_{33} \end{bmatrix} \text{ and } \bar{\mathbf{C}} = \begin{bmatrix} \bar{\omega}_{11} & \bar{\omega}_{12} & \bar{\omega}_{13} \\ \bar{\omega}_{21} & \bar{\omega}_{22} & \bar{\omega}_{23} \\ \bar{\omega}_{31} & \bar{\omega}_{32} & \bar{\omega}_{33} \end{bmatrix},$$

and from (7) we have

$$\begin{bmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} & \hat{\omega}_{13} \\ \hat{\omega}_{21} & \hat{\omega}_{22} & \hat{\omega}_{23} \\ \hat{\omega}_{31} & \hat{\omega}_{32} & \hat{\omega}_{33} \end{bmatrix} = \begin{bmatrix} \bar{\omega}_{11}f^2 & \bar{\omega}_{12}f^2 & \bar{\omega}_{13}f \\ \bar{\omega}_{21}f^2 & \bar{\omega}_{22}f^2 & \bar{\omega}_{23}f \\ \bar{\omega}_{31}f & \bar{\omega}_{32}f & \bar{\omega}_{33} \end{bmatrix} \tag{8}$$

Recall that $\hat{\mathbf{C}}$ can be factorized into $\hat{\mathbf{C}} = \mathbf{R}\mathbf{D}\mathbf{R}^{\mathrm{T}}$, where $\mathbf{D} = diag(\lambda_1, \lambda_1, \lambda_2)$ is a diagonal matrix composed of the

eigenvalues of $\hat{\mathbf{C}}$. The eigenvalues of $\hat{\mathbf{C}}$ can be obtained by solving the characteristic equation

$$det(\hat{\mathbf{C}} - \lambda\mathbf{I}) = 0. \tag{9}$$

Note that (9) gives a cubic polynomial

$$\lambda^3 - \beta\lambda^2 + \gamma\lambda + \delta = 0, \tag{10}$$

where

$$\beta = \hat{\omega}_{11} + \hat{\omega}_{22} + \hat{\omega}_{33}$$
$$\gamma = \hat{\omega}_{11}\hat{\omega}_{22} + \hat{\omega}_{11}\hat{\omega}_{33} + \hat{\omega}_{22}\hat{\omega}_{33} - \hat{\omega}_{12}^2 - \hat{\omega}_{13}^2 - \hat{\omega}_{23}^2$$
$$\delta = \hat{\omega}_{11}\hat{\omega}_{23}^2 + \hat{\omega}_{22}\hat{\omega}_{13}^2 + \hat{\omega}_{33}\hat{\omega}_{12}^2 - \hat{\omega}_{11}\hat{\omega}_{22}\hat{\omega}_{33}$$
$$- 2\hat{\omega}_{12}\hat{\omega}_{13}\hat{\omega}_{23}$$

Since $\hat{\mathbf{C}}$ has at least two identical eigenvalues, (10) must have at least two equal roots. Hence we have [6]

$$\mu^2 - 4\nu^3 = 0, \tag{11}$$

where

$$\mu = \beta^2 - 3\gamma$$
$$= \frac{1}{2}[(\hat{\omega}_{11} - \hat{\omega}_{22})^2 + (\hat{\omega}_{11} - \hat{\omega}_{33})^2 + (\hat{\omega}_{22} - \hat{\omega}_{33})^2]$$
$$+ 3(\hat{\omega}_{12}^2 + \hat{\omega}_{13}^2 + \hat{\omega}_{23}^2)$$
$$\nu = 2\beta^2 - 9\beta\gamma - 27\delta$$
$$= 18(\hat{\omega}_{11}\hat{\omega}_{22}\hat{\omega}_{33} + 3\hat{\omega}_{12}\hat{\omega}_{13}\hat{\omega}_{23})$$
$$+ 2(\hat{\omega}_{11}^2 + \hat{\omega}_{22}^2 + \hat{\omega}_{33}^2)$$
$$+ 9(\hat{\omega}_{11} + \hat{\omega}_{22} + \hat{\omega}_{33})(\hat{\omega}_{12}^2 + \hat{\omega}_{13}^2 + \hat{\omega}_{23}^2)$$
$$- 3(\hat{\omega}_{11} + \hat{\omega}_{22})(\hat{\omega}_{11} + \hat{\omega}_{33})(\hat{\omega}_{22} + \hat{\omega}_{33})$$
$$- 27(\hat{\omega}_{11}\hat{\omega}_{23}^2 + \hat{\omega}_{22}\hat{\omega}_{13}^2 + \hat{\omega}_{33}\hat{\omega}_{12}^2)$$

Given the conic $\mathbf{C}$ and the principal point (u_0, v_0), the only unknown in (11) is f. Solving (11) leads to 12 solutions. We observe that among these 12 solutions, 8 of them are zeros, and the remaining 4 have an identical absolute value, which gives the solution of f. Under noisy data, the solutions to (11) might be complex numbers. The 8 zero solutions become pure imaginary numbers, and the remaining 4 become two pairs of conjugate complex numbers, with opposite signs for their real and imaginary parts (i.e., $a + bi$, $a - bi$, $-a + bi$, $-a - bi$). The absolute value of their real parts gives us an estimate of the focal length, i.e., $f = \|a\|$.

4.2. Principal Point

As discussed above, a unique focal length f can be obtained by solving (11) when the principal point of the camera is known or assumed to locate at the image centre. In practice, however, the principal point is often not known a priori. Besides, due to manufacture reasons, the principal

point usually does not locate exactly at the image centre, but somewhere close to it.

We notice that given the ground truth principal point (u_0, v_0), the focal length estimated by (11) should be identical to the ground truth focal length f_{gt}. On the other hand, if we solve (11) using points close to the ground truth principal point as the principal point, the estimated focal length will become a complex number. The real part of this complex number is close to f_{gt}, and its imaginary part is a relatively small value to compensate the errors.

Another observation is that, given the conic images of the mirror sphere at two distinct positions, we can obtain two estimates of the focal length, denoted as f_1 and f_2 respectively, using the same assumed principal point. $f_1 = f_2 = f_{gt}$ holds when the assumed principal point equals the ground truth principal point.

Based on the above observations, we propose to estimate the principal point by minimizing the difference between the focal lengths estimated from the conic images of the mirror sphere at two distinct positions, subject to the principal point lying within a small window centred at the image centre. The difference between the estimated focal lengths is measured as the sum of the magnitudes of the differences in their real and imaginary parts respectively

$$error = \|real(f_1)| - |real(f_2)\|$$
$$+ \|imag(f_1)| - |imag(f_2)\|. \tag{12}$$

The proposed approach is summarized in Algorithm 1.

Algorithm 1: Estimation of the principal point and focal length from two conic images of a mirror sphere.

Input : Image centre (u_c, v_c), conic images $\mathbf{C}_1$, $\mathbf{C}_2$
Output: Principal point (u_0, v_0), focal length f
Initialization: Set offsets along u, v directions as $w = const1$, $h = const2$; set step size as $s = const3$;
set error as $error_{min} = large\ const$;
for $u_p \leftarrow u_c - w$ **to** $u_c + w$ **step** s **do**
  for $v_p \leftarrow v_c - h$ **to** $v_c + h$ **step** s **do**
    Construct $\mathbf{T}$ using (4) with (u_p, v_p);
    Compute $\bar{\mathbf{C}}_1$, $\bar{\mathbf{C}}_2$ using (6);
    Construct $\hat{\mathbf{C}}_1$, $\hat{\mathbf{C}}_2$ from $\bar{\mathbf{C}}_1$, $\bar{\mathbf{C}}_2$ using (7);
    Solve (11) for $\hat{\mathbf{C}}_1$, $\hat{\mathbf{C}}_2$ to obtain f_1, f_2;
    Compute current error, $error_{cur}$, by (12);
    if $error_{cur} < error_{min}$ **then**
      $u_0 \leftarrow u_p$
      $v_0 \leftarrow v_p$
      $f \leftarrow \frac{|real(f_1)| + |real(f_2)|}{2}$
      $error_{min} \leftarrow error_{cur}$
    end
  end
end

The above algorithm requires two conic images of the mirror sphere. When only one image of the sphere is available, one simple strategy is to assume the principal point is at the image centre and then estimate f by solving (11). This, however, often does not give an optimal solution as the principal point, as mentioned before, usually does not locate exactly at the image centre. Based on the observation that the error in the estimated focal length is highly correlated with the error in the position of the principal point, we propose a novel approach for estimating both the principal point and focal length of the camera from just one single image of the mirror sphere. We first sample points evenly within a small window centred at the image centre, and estimate a focal length f using each sample point as the principal point. We then calculate the mean of these estimated values and this gives us a final estimate of the focal length. Next, we identify the sample point that leads to an estimated focal length closest to the mean value as the principal point. Fig. 3 gives an example of estimating the focal length using points sampled around the image centre as the principal point. We can see that when the sampled principal point is closer to the ground truth principal point, the estimated f will be closer to the ground truth focal length f_{gt}.
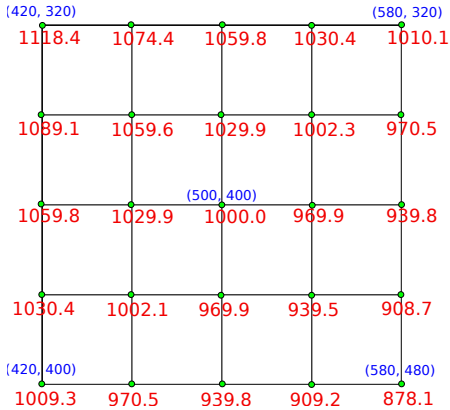

Figure 3. An example of camera intrinsics estimation from the conic of a single sphere. The ground truth intrinsic parameters are $f = 1000$, $(u_0, v_0) = (500, 400)$. Given the conic of the sphere, we estimated the focal length by setting different image points as the principal point. The coordinates (blue) stand for the pixel coordinates, and the number (red) close to a pixel stands for the estimated f by setting that pixel as the principal point. The mean of the estimated values is $f_m = 998.1$. $(500, 400)$ is the sample point that leads that to an estimated f closest to f_m.

$\mathbf{C}$ into $\hat{\mathbf{C}}$ using (1), and factorize it into $\hat{\mathbf{C}} = \mathbf{RDR}^{\mathrm{T}}$ by eigen decomposition. The centre of the sphere, expressed in camera coordinates, can be recovered as $d\mathbf{r}_3$ where $\mathbf{r}_3$ denotes the third column of $\mathbf{R}$ resulting from the eigen decomposition, and $d = r_s\sqrt{\frac{1+r_c^2}{r_c}}$ denotes the distance of the sphere centre from the camera centre. With the radius r_s of the sphere being unknown, we can simply set r_s to 1 and the resulting reconstruction will then be scaled by $\frac{1}{r_s}$.

Consider a scene point $\mathbf{P}$, and let $\mathbf{q}_1$ and $\mathbf{q}_2$ be its reflections observed on the surface of a mirror sphere placed at $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively (see Fig. 4). To reconstruct $\mathbf{P}$ (in camera coordinates), we first construct the visual rays $\mathbf{V}(\mathbf{q}_1)$ and $\mathbf{V}(\mathbf{q}_2)$ for $\mathbf{q}_1$ and $\mathbf{q}_2$, respectively, using the formula $\mathbf{V}(\mathbf{q}) = \mathbf{K}^{-1}\mathbf{q}$. We solve for the point of intersection $\mathbf{Q}_1$ between $\mathbf{V}(\mathbf{q}_1)$ and the sphere at $\mathbf{S}_1$, and the point of intersection $\mathbf{Q}_2$ between $\mathbf{V}(\mathbf{q}_2)$ and the sphere at $\mathbf{S}_2$, respectively. Based on the law of reflection, the incident rays at $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can be constructed, and $\mathbf{P}$ can be recovered by triangulating these incident rays. In the case where $\mathbf{P}$ can be directly observed by the camera, only one reflection of $\mathbf{P}$ on the surface of a mirror sphere is sufficient to recover $\mathbf{P}$ by triangulating the visual ray of the direct observation with the incident ray of the reflection observed.
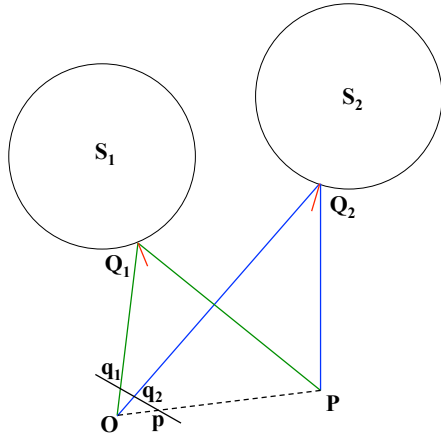

Figure 4. A perspective camera located at $\mathbf{O}$ observes the reflections of a scene point $\mathbf{P}$ at $\mathbf{Q_1}$ and $\mathbf{Q_2}$ on the surface of a mirror sphere placed at $\mathbf{S_1}$ and $\mathbf{S_2}$, respectively. $\mathbf{P}$ can be reconstructed by triangulating: (a) $\mathbf{Q_1}\mathbf{P}$ with $\mathbf{Q_2}\mathbf{P}$, if $\mathbf{P}$ is not visible; (b) $\mathbf{OP}$ with $\mathbf{Q_1}\mathbf{P}$ (or $\mathbf{Q_2}\mathbf{P}$), if $\mathbf{P}$ is visible.

5. Shape Recovery

After estimating the principal point (u_0, v_0) and focal length f of the camera using the methods described in the Section 4, we can obtain the camera calibration matrix $\mathbf{K}$ using (3) and (4). We use $\mathbf{K}$ to transform the conic image

6. Experimental Results

We evaluate our proposed methods on both synthetic and real data. We compare our uncalibrated approach against the method in [7] which assumes an internally calibration camera.
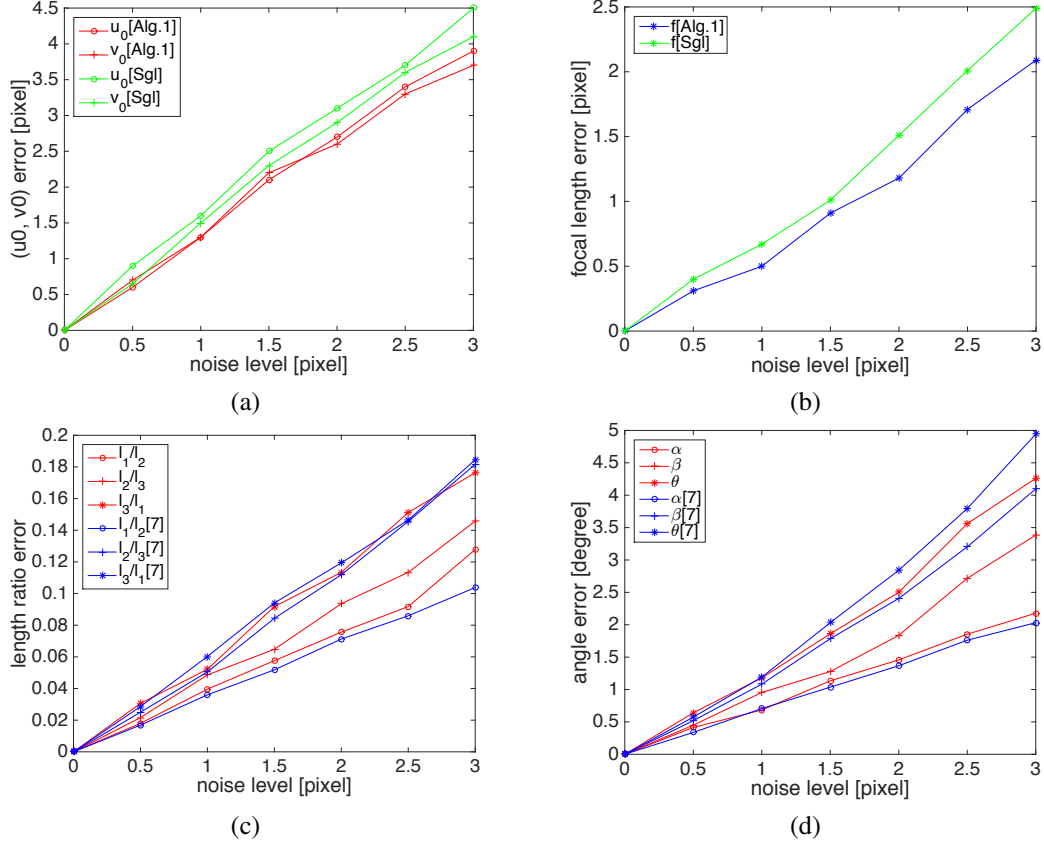
Figure 5. Synthetic experiment results under different noise levels. Ground truth value: $f = 1200$, $u_0 = 495$, $v_0 = 395$. (a) Reconstructed principal point (u_0, v_0) errors by Algorithm 1 **[Alg.1]**, and by the method using only a single image of the sphere **[Sgl]**. (b) Focal length errors; (c) Length ratio errors compared with [7]; (d) Angle errors of the reconstructed rays compared with [7].

6.1. Synthetic Data

We used the same synthetic data set as in [7] to evaluate our approach and compare against their method. A synthetic perspective camera was employed to observe a mirror sphere with radius $1.5mm$ placed at two distinct positions. Four 3D points $\mathbf{P}_{\{1,2,3,4\}}$ were reflected at $\mathbf{Q}_{\{11,12,13,14\}}$ on the surface of the sphere when it was placed at $\mathbf{S}_1$, and at $\mathbf{Q}_{\{21,22,23,24\}}$ on the surface of the sphere when it was placed at $\mathbf{S}_2$. $\mathbf{P}_1\mathbf{P}_2$, $\mathbf{P}_1\mathbf{P}_3$ and $\mathbf{P}_1\mathbf{P}_4$ were mutually orthogonal to each other, and they were of the same lengths ($5\ mm$).

We applied SVD to fit conics to the edges extracted from the images of the sphere [12]. To evaluate the robustness of our approach, we added Gaussian noise to the pixel coordinates of the edges with noise level ranging from 0 to 3.0 pixels. With conics fitted to two images of the mirror sphere, the camera calibration matrix $\mathbf{K}$ can be estimated using our method described in Section 4.

With an unknown radius of the sphere, we can only reconstruct the scene up to an unknown scale. Hence, it is more meaningful to compare the reconstruction error in
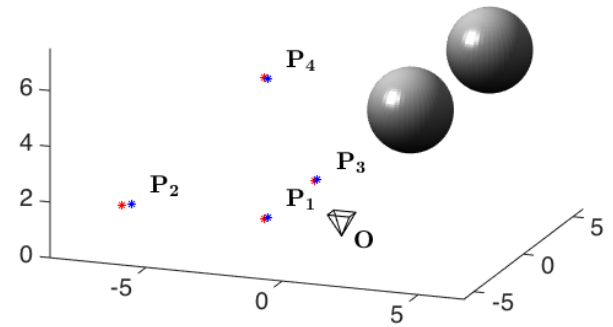

Figure 6. Reconstruction under noise level $\sigma = 2.0$. Blue points: ground truth; Red points: reconstruction. (Best viewed in color.)

terms of length ratios and angles rather than absolute distance errors. Let l_1, l_2, and l_3 denote $\|\mathbf{P}_1\mathbf{P}_2\|$, $\|\mathbf{P}_1\mathbf{P}_3\|$ and $\|\mathbf{P}_1\mathbf{P}_4\|$, respectively. We measured the errors in the length ratios l_1/l_2, l_2/l_3 and l_3/l_1, respectively, as their deviations from 1. Similarly, let α, β and θ denote $\angle\mathbf{P}_2\mathbf{P}_1\mathbf{P}_3$, $\angle\mathbf{P}_3\mathbf{P}_1\mathbf{P}_4$ and $\angle\mathbf{P}_4\mathbf{P}_1\mathbf{P}_2$, respectively. The errors in α, β and θ were measured as their deviations from $90°$. We

performed 500 independent trials and our reconstruction results as well as the comparison with [7] are presented in Fig. 5. Generally, the errors increase linearly with the noise level. Algorithm 1 performed slightly better than the method using only a single image of the sphere, hence we used the intrinsics estimated by Algorithm 1 for 3D reconstruction in our experiment. Our length ratio errors and angle errors are quite close to that of [7]. Fig. 6 shows our reconstruction under the noise level $\sigma = 2.0$. The results demonstrate the accuracy and robustness of our proposed approach.

6.2. Real Data

To evaluate our approach on real data set, we performed an experiment on a rectangular box with a dimension of $6\ cm \times 6\ cm \times 8\ cm$. In the experiment, the box was reflected by a mirror sphere of radius $40\ mm$ placed at four different positions. The reflections on the sphere surface were captured using a $Canon\ EOS\ 40D$ equipped with a 24-70 mm lens. Fig. 7 shows our experiment setup.
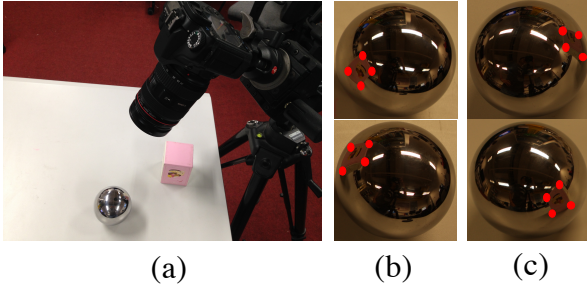

| | (a) | | (b) | (c) |

Figure 7. Real experiment setup (a) and reflections on a sphere at four distinct positions (b-c). Four corners of a box are reconstructed using the reflections in (b) and the other four corners are reconstructed using the reflections in (c). The correspondences are marked with red dots in the images. Note that the box is not visible by the camera.

After fitting conics to the images of the sphere using SVD, we first estimated the principal point and focal length of the camera using the methods introduced in Section 4. Table 1 shows our estimation result. It can be seen that the estimated camera intrinsic parameters are very close to those obtained by camera calibration using a calibration pattern [5]. We used the intrinsics estimated by Algorithm 1 to reconstruct the corners of the box, and measured 24 angles and 12 length ratios around them. We compared these measurements against the ground truth values ($90°$ for angle and 1.25 for length ratio). We also compared our results with that of [7], which works under an internally calibrated camera. The RMS errors are given in Table 2. Fig. 8 shows the reconstructed 3D corner points of our method and that of [7]. The recovered surfaces are presented in Fig. 9. Our

approach achieved a high accuracy which is very close to that using a calibrated camera.

	f	u_0	v_0
$Calibration$ [5]	4435.36	1963.0	1277.0
$Estimation$ [Centre]	4117.49	1944.0	1296.0
$Estimation$ [Alg.1]	4386.06	1955.0	1285.0
$Estimation$ [Sgl]	4301.02	1980.0	1290.0
$Error$ [Centre]	7.17%	0.97%	1.49%
$Error$ [Alg.1]	**1.11%**	**0.41%**	**0.63%**
$Error$ [Sgl]	3.06%	0.87%	1.02%

Table 1. Estimation of camera intrinsic parameters. **[Centre]**: results by setting image centre as the principal point; **[Alg.1]**: results by Algorithm 1; **[Sgl]**: results by the method using only a single image of the sphere.

	angle	length ratio
$Calibrated$[7]	1.08	0.03
$Ours$	1.05	0.03

Table 2. RMS angle error and length ratio error of the recovered box. The ground truth angle and length ratio are $90°$ and 1.25, respectively.
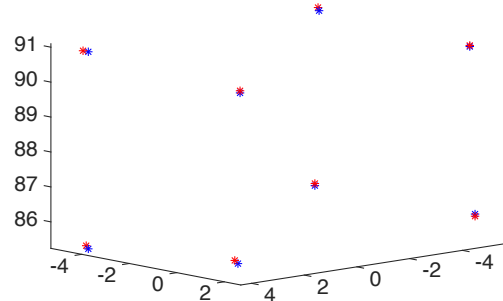

Figure 8. Reconstructed corners of a box. Blue: using a calibrated camera [7]; Red: using an uncalibrated camera (ours). (Best viewed in color.)
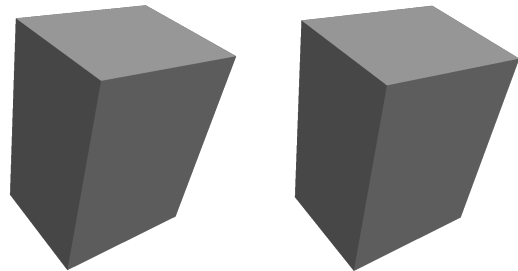

Figure 9. Recovered surfaces. Left: using a calibrated camera [7]; Right: using an uncalibrated camera (ours).

7. Discussions and Conclusions

This paper addresses the problem of single view 3D reconstruction under an uncalibrated camera and an unknown mirror sphere. We derive an analytical solution to solve the camera focal length given the principal point by enforcing a repeated eigenvalue constraint for the conic image of the mirror sphere. Based on this analytical solution , we introduce a robust algorithm to estimate both the principal point and focal length of the camera by minimizing the difference between focal lengths estimated from multiple images of the sphere. Besides, we also present a novel approach to estimate both the principal point and focal length of the camera in the case of just one single image of the sphere. With the estimated camera intrinsic parameters, we can recover the sphere position by eigen decomposition, and reconstruct the scene up to an unknown scale determined by the radius of the sphere. Experimental results on both synthetic and real data demonstrate the feasibility and accuracy of our proposed approach. One limitation of our approach is the difficulty in establishing dense and high quality reflection correspondences due to the distortions exhibited in the reflected images. As a result, the reconstructed point cloud is sparse. Note that this is in fact the limitation of all 3D reconstruction methods based on mirror spheres. In the future, we would like to extend our work to achieve dense 3D reconstruction.

References

[1] A. Agrawal. Extrinsic camera calibration without a direct view using spherical mirror. In *ICCV*, pages 2368–2375, 2013. 2

[2] A. Agrawal and S. Ramalingam. Single image calibration of multi-axial imaging systems. In *CVPR*, pages 1399–1406, 2013. 2

[3] M. Agrawal and L. S. Davis. Camera calibration using spheres: A semi-definite programming approach. In *ICCV*, pages 782–789, 2003. 2

[4] J. Balzer and S. Werling. Principles of shape from specular reflection. *Measurement*, 43(10):1305–1317, 2010. 2

[5] J.-Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/. 7

[6] G. Cardano. *The Rules of Algebra (Ars Magna)*. Dover Publications, 2007. 4

[7] Z. Chen, K.-Y. K. Wong, M. Liu, and D. Schnieders. Single-view reconstruction from an unknown spherical mirror. In *ICIP*, pages 2733–2736, 2011. 1, 2, 5, 6, 7

[8] Y. Ding, J. Yu, and P. Sturm. Multi-perspective stereo matching and volumetric reconstruction. In *ICCV*, pages 1827–1834, 2009. 1, 2

[9] J. Gluckman and S. K. Nayar. Planar catadioptric stereo: Geometry and calibration. In *CVPR*, volume 1, pages 22–28, 1999. 1, 2

[10] J. Gluckman and S. K. Nayar. Catadioptric stereo using planar mirrors. *IJCV*, 44(1):65–79, 2001. 1, 2

[11] J. Gluckman and S. K. Nayar. Rectified catadioptric stereo sensors. *PAMI*, 24(2):224–236, 2002. 1, 2

[12] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004. 1, 2, 6

[13] B. Hu. Its all done with mirrors: Calibration-and-correspondence-free 3d reconstruction. In *CVR*, pages 148–154, 2009. 2

[14] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich. State of the art in transparent and specular object reconstruction. In *Eurographics STAR*, pages 87–108, 2008. 2

[15] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich. Transparent and specular object reconstruction. *Computer Graphics Forum*, 29:2400–2426, 2010. 2

[16] I. Ihrke, T. Stich, H. Gottschlich, M. Magnor, and H. Seidel. Fast incident light field acquisition and rendering. In *WSCG*, volume 16, pages 25–32, 2008. 2

[17] G. Jang, S. Kim, and I. Kweon. Single camera catadioptric stereo system. In *OMNIVIS*, 2005. 1

[18] R. Jensen, A. Dahl, G. Vogiatzis, E. Tola, and H. Aans. Large scale multi-view stereopsis evaluation. In *CVPR*, pages 406–413, 2014. 1

[19] M. Kanbara, N. Ukita, M. Kidode, and N. Yokoya. 3d scene reconstruction from reflection images in a spherical mirror. In *ICPR*, pages 874–879, 2006. 2

[20] D. Lanman, D. Crispell, M. Wachs, and G. Taubin. Spherical catadioptric arrays: Construction, multiview geometry, and calibration. In *3DPVT*, pages 81–88, 2006. 1, 2

[21] D. Lanman, M. Wachs, G. Taubin, and F. Cukierman. Reconstructing a 3d line from a single catadioptric image. In *3DPVT*, pages 1–8, 2006. 1

[22] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-d reconstruction using mirror images based on a plane symmetry recovering method. *PAMI*, 14:941–947, 1992. 2

[23] S. Nayar. Sphereo: Determining depth using two specular spheres and a single camera. In *SPIE Conf. Optics, Illumination, and Image Sensing for Machine Vision III*, pages 245–254, 1988. 2

[24] C. Nitschke, A. Nakazawa, and H. Takemura. Display-camera calibration using eye reflections and geometry constraints. *CVIU*, 115:835–853, 2011. 2

[25] L. Perdigoto and H. Araujo. Calibration of mirror position and extrinsic parameters in axial non-central catadioptric systems. *CVIU*, 117:909–921, 2013. 2

[26] M. Powell, S. Sarkar, and D. Goldgof. A simple strategy for calibrating the geometry of light sources. *PAMI*, 23:1022–1027, 2001. 2

[27] I. Reshetouski and I. Ihrke. Mirrors in computer graphics, computer vision and time-of-flight imaging. *Lect. Notes Comput. Sc.*, 8200:77–104, 2013. 2

[28] R. Sagawa, N. Kurita, T. Echigo, and T. Yagi. Compound catadioptric stereo sensor for omnidirectional object detection. In *IROS*, pages 2612–2617, 2004. 2

[29] D. Schnieders, X. Fu, and K.-Y. K. Wong. Reconstruction of display and eyes from a single image. In *CVPR*, pages 1442–1449, 2010. 2

[30] D. Schnieders and K.-Y. K. Wong. Camera and light calibration from reflections on a sphere. *CVIU*, 117:1536–1547, 2013. 2

[31] P. Sturm and T. Bonfort. How to compute the pose of an object without a direct view. In *ACCV*, pages 21–31, 2006. 2

[32] J. Sun, X. Chen, Z. Gong, Z. Liu, and Y. Zhao. Accurate camera calibration with distortion models using sphere images. *Optics & Laser Technology*, 65:83–87, 2015. 2

[33] Y. Taguchi, A. Agrawal, A. Veeraraghavan, S. Ramalingam, and R. Raskar. Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering. *ACM Transactions on Graphics (TOG)*, 29(6):172:1–172:8, 2010. 2

[34] K. Tan, H. Hua, and N. Ahuja. Multiview panoramic cameras using mirror pyramids. *PAMI*, 26(7):941–946, 2004. 2

[35] K.-Y. K. Wong, D. Schnieders, and S. Li. Recovering light directions and camera poses from a single sphere. In *ECCV*, volume I, pages 631–642, 2008. 2, 3

[36] K.-Y. K. Wong, G. Zhang, and Z. Chen. A stratified approach for camera calibration using spheres. *TIP*, 20(2):305–316, 2011. 2

[37] X. Ying and H. Zha. Geometric interpretations of the relation between the image of the absolute conic and sphere images. *PAMI*, 28(12):2031–2036, 2006. 2

[38] H. Zhang, K.-Y. K. Wong, and G. Zhang. Camera calibration from images of spheres. *PAMI*, 29(3):499–503, 2007. 2

[39] H. Zhong, W. F. Sze, and Y. S. Hung. Reconstruction from plane mirror reflection. In *ICPR*, pages 715–718, 2006. 2